Structural Equivalence and Assortative Mixing

Leonid E. Zhukov

National Research University Higher School of Economics

27.02.2012

Structural Equivalence

Two vertices are structurally equivalent if they share the same neighbours

Number of shared neighbours ($\mathbf{A_{ij}}$ - binary adjacency matrix):

$$n_{ij} = \sum_k A_{ik} A_{kj} = (A^2)_{ij}$$

Euclidean distance (Hamming distance):

$$d_{ij}^2 = \sum_k (A_{ik} - A_{jk})^2$$

Maximal possible:

$$max(d_{ij}^2) = k_i + k_j$$

Normalized Hamming distance:

$$d_{ij}^2{}_N = 1 - \frac{2n_{ij}}{k_i + k_j}$$

Structural Equivalence

Similarity (vector)

$$\sigma_{ij} = cos(\theta_{ij}) = \frac{xy}{|x||y|} = \frac{\sum_k A_{ik}A_{kj}}{\sqrt{\sum A_{ik}A_{ki}}\sqrt{\sum A_{jk}A_{kj}}} = \frac{n_{ij}}{\sqrt{k_i k_j}}$$

Correlation coefficient:

$$r_{ij} = \frac{\sum_k (A_{ik} - \langle A_i \rangle)(A_{jk} - \langle A_j \rangle)}{\sqrt{\sum_k (A_{ik} - \langle A_i \rangle)^2}\sqrt{\sum_k (A_{jk} - \langle A_j \rangle)^2}}$$

$A_{ik} = A_{ki}, \quad \sum_k A_{ik}^2 = \sum_k A_{ik} = k_i$
$\langle A_i \rangle = \frac{1}{n} \sum_k A_{ik}$

Assortative Mixing

Assortative mixing (Homophily)
Tendency to associate and form connections with those percieved to be similar.
Let n_c classes, c_i class, $\delta(c_i, c_j)$ kronecker delta
Number of edges between nodes of the same class in

- given network A_{ij} $m_c = \sum \delta(c_i, c_j) = \frac{1}{2} \sum_{ij} A_{ij} \delta(c_i, c_j)$
- random network $\langle m_c \rangle = \frac{1}{2} \sum_{ij} \frac{k_i k_j}{2m} \delta(c_i, c_j)$

Modularity:

$$Q = \frac{1}{2m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(c_i, c_j)$$

Assortative Mixing

Assortative mixing: every node has value x_i
Average and covariance over edges

$$\langle x \rangle = \frac{1}{2m} \sum_{ij} A_{ij} x_i = \frac{1}{2m} \sum_i k_i x_i$$

$$var = \frac{1}{2m} \sum_{ij} A_{ij} (x_i - \langle x \rangle)^2 = \frac{1}{2m} \sum_i k_i (x_i - \langle x \rangle)^2$$

$$cov = \frac{1}{2m} \sum_{ij} A_{ij} (x_i - \langle x \rangle)(x_j - \langle x \rangle)$$

Assortativity coefficient

$$r = \frac{cov}{var} = \frac{\sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) x_i x_j}{\sum_{ij} \left(k_i \delta_{ij} - \frac{k_i k_j}{2m} \right) x_i x_j}$$

Assortative Mixing

Assortative mixing by degree

$$r = \frac{\sum_{ij}\left(A_{ij} - \frac{k_i k_j}{2m}\right) k_i k_j}{\sum_{ij}\left(k_i \delta_{ij} - \frac{k_i k_j}{2m}\right) k_i k_j}$$

$S_1 = \sum_i k_i = 2m$
$S_2 = \sum_i k_i^2$
$S_3 = \sum_i k_i^3$
$S_e = \sum_{ij} A_{ij} k_i k_j = \sum_{e \in E} k_i k_j$

$$r = \frac{S_e S_1 - S_2^2}{S_3 S_1 - S_2^2}$$